GENES CODING FOR TOMATO β-GALACTOSIDASE

POLYPEPTIDES

5

Field of the Invention

The present invention relates to a family of novel plant genes encoding

polypeptides characterized by their ability to hydrolyze terminal non-reducing

10   β-D-galactosyl residues from β-D-galactosides. More specifically, a

polynucleotide sequence derived from a cDNA clone designated pZBG2-1-4

(referred to in U.S. Provisional Appln. No. 60/088,805 as pTomβgal 4), which

encodes a specific plant polypeptide named β-galactosidase II, is provided.

Also provided are cDNA clones encoding six other homologous polypeptides,

15   methods of using these cDNA clones for producing β-D-galactoside

polypeptides of the invention, and methods of modifying fruit quality by

employment of a polynucleotide or polypeptide of the present invention.

Background of the Invention

The most conspicuous and important processes related to post-harvest

20   quality of climacteric fruit are the changes in texture, color, taste, and aroma

which occur during ripening. Because of the critical relationship that

deleterious changes in texture have to quality and post-harvest shelf-life,

emphasis has been placed on studying the mechanisms involved in the loss of

firmness that occurs during tomato fruit ripening. Although fruit softening

25   may involve changes in turgor pressure, anatomical characteristics and cell

wall integrity, it is generally assumed that cell wall disassembly leading to a

loss of wall integrity is a critical feature. The most apparent changes, in terms

of composition and size, occur in the pectic fraction of the cell wall (see

references in Seymour and Gross, 1996).

5   Changes known to occur in the pectic fraction of the cell wall during fruit

ripening include increased solubility, depolymerization, de-esterification and a

significant net loss of neutral sugar containing side chains (Huber, 1983; ·

Fischer and Bennett, 1991; Seymour and Gross, 1996). The best characterized

pectin-modifying enzymes are polygalacturonase (endo-$\alpha1\rightarrow$4-D-galacturonan

10   hydrolase; E.C. 3.2.1.15; PG) and pectin methylesterase (E.C. 3.1.1.11; PME).

Although PG and PME are relatively abundant and have substantial activity

during tomato fruit ripening, softening still occurs, albeit with a slight delay, in

fruit where PG (Smith *et al.* 1988, 1990) or PME (Tieman *et al.* 1992; Hall *et*

al. 1993) gene expression and enzyme activity was significantly down-

15   regulated in transgenic plants. Moreover, over-expression of PG in non-

ripening mutant *rin* tomato fruit did not result in softening even though

depolymerization and solubilization of pectin was evident (Giovannoni *et al.*,

1989).

Among the other known pectin modifications that occur during fruit

20   development, one of the best characterized is the significant net loss of

galactosyl residues which occurs in the cell walls of many ripening fruit (Gross

and Sams, 1984; Seymour and Gross, 1996). Although some loss of galactosyl

residues could result indirectly from the action of PG, β-galactosidase (exo-

β(1$\rightarrow$4)-D-galactopyranoside; E.C. 3.2.1.23) is the only enzyme identified in

higher plants capable of directly cleaving $\beta(1\rightarrow4)$galactan bonds, and probably

plays a role in galactan sidechain loss (DeVeau et al., 1993; Carey et al., 1995;

Carrington and Pressey, 1996). No endo-acting galactanase has yet been

identified in higher plants. The view that β-galactosidase is active in releasing

galactosyl residues from the cell wall during ripening is supported by the

dramatic increase in free galactose, a product of β-galactosidase activity

(Gross, 1984) and a concomitant increase in activity of a particular enzyme,

designated β-galactosidase II, in tomatoes during ripening (Carey et al., 1995).

β-galactosidase activity is thought to be important in cell wall metabolism

(Carey et al., 1995). β-Galactosidases are generally assayed using artificial

substrates such as p-nitrophenyl-β-D-galactopyranoside (PNP), 4-

methylumbelliferyl-β-D-galactopyranoside and 5-bromo-4-chloro-3-indoxyl-

β-D-galactopyranoside (X-GAL). However, it is clear that β-galactosidase II

is also active against natural substrates, i.e., β $(1\rightarrow4)$galactan (Carey et al.,

1995; Carrington and Pressey, 1996; Pressey, 1983). β-Galactosidase proteins

have been purified and characterized in a number of other fruits including

kiwifruits (Ross et al., 1993), coffee (Golden et al., 1993), persimmon (Kang

et al., 1994), and apple (Ross et al., 1994).

Carey et al. (1995) were able to purify three previously identified β-

galactosidases from ripening tomato fruit (Pressey, 1983), but only one (β-

galactosidase II) was active against $\beta(1\rightarrow4)$galactan. Even though they were

able to identify putative β-galactosidase cDNA clones, none of the cDNA's

deduced amino acid sequences matched the amino terminal sequence of the β-

galactosidase II protein. Although β-galactosidase II, a protein present in

tomato (*Lycopersicon esculentum* Mill.) fruit during ripening and capable of

degrading tomato fruit galactan has been purified, cloning of the

corresponding gene has been elusive.

The modification of plant gene expression has been achieved by several

5   methods. The molecular biologist can choose from a range of known methods

to decrease or increase gene expression or to alter the spatial or temporal

expression of a particular gene. For example, the expression of either specific

antisense RNA or partial (truncated) sense RNA has been utilized to reduce

the expression of various target genes in plants (as reviewed by Bird and Ray,

10   1991, Biotechnology and Genetic-Engineering Reviews 9:207-227). These

techniques involve the incorporation into the genome of the plant of a

synthetic gene designed to express either antisense or sense RNA. They have

been successfully used to down-regulate the expression of a range of

individual genes involved in the development and ripening of tomato fruit

15   (Gray et al, 1992, Plant Molecular Biology, i9:69-87). Methods to increase the

expression of a target gene·have also been developed. For example, additional

genes designed to express RNA containing the complete coding region of the

target gene may be incorporated into the genome of the plant to "over-express"

the gene product. Various other methods to modify gene expression are

20   known; for example, the use of alternative regulatory sequences. The

complete disclosure of each of the references cited above is fully incorporated

herein by reference.

The need therefore exists to clone a gene for β-galactosidase II and related

polypeptides, and using known methods of modification of plant gene

25   expression, thereby to provide methods for modifying quality of fruits,

particularly by modifying the cell wall, thereby directly affecting the ripening

of the fruit.

Summary of the Invention

The present invention is based on the discovery of novel DNA sequences

derived from cDNA clones from a family of genes encoding β-galactosidases.

The phylogenic tree based on the shared amino acid sequence identities for the

DNA sequences of the present invention is shown in Figure 1A,B. Five cDNA

and two RT-PCR clones, designated herein as TBG1, TBG2, TBG3, TBG4,

TBG5, TBG6, and TBG7 and having the nucleic acid sequences designated

SEQ ID NOs 1-7, respectively as shown in Figure 2, were identified which had

a high degree of shared sequence identity to other known β-galactosidases.

The corresponding amino acid sequences are designated herein as SEQ ID

NOs 8-16, respectively and are shown in Figure 2 and 3. The nucleotide

sequences for SEQ ID NOs 1-7 are recorded in Gen Bank with the following

respective Accessions Numbers:

SEQ ID NO:1	TGB1	AF023847	deposit Sept 10, 1997
SEQ ID NO:2	TGB2	AF154420	deposited May 19, 1999
SEQ ID NO: 3	TGB3	AF154421	deposited May 20, 1999
SEQ ID NO:4	TGB4	AF020390	deposited Aug 21, 1997
SEQ ID NO:5	TGB5	AF154423	deposited May 20, 1999
SEQ ID NO:6	TGB6	AF154424	deposited May 20, 1999
SEQ ID NO: 7	TGB7	AF154422	deposited May 20, 1999

Throughout the following discussion, wherever TBG4 is indicated in the description of the invention, it is to be understood that TBG1-3 and 5-7 are also to be included in that description, unless otherwise indicated.

A method of providing a DNA sequence of the invention, either by cloning a cDNA (for instance, pZBG2-1-4) that codes for a protein of the present invention, such as β-galactosidase II, or by deriving the DNA sequence from genomic DNA, or by synthesis of a DNA sequence ab initio using the cDNA sequence as a guide is also provided.

A method for modifying cell wall metabolism which involves modifying the activity of at least one galactosidase, and thus modifying the quality of the fruit is also provided.

Also provided by the present invention is a DNA construct including some or all of an exemplary β-galactosidase DNA sequence under control of a transcriptional initiation region operative in plants, so that the construct can generate RNA in plant cells.

Also discovered is an enhancer/promoter associated with expression of the genes encoding β-galactosidase.

The present invention also relates to recombinant vectors, which include the isolated nucleic acid molecules of the present invention, and to host cells containing the recombinant vectors, as well as to methods of making such vectors and host cells and for using them for production of β-galactosidase polypeptides or peptides by recombinant techniques.

The present invention also provides plant cells containing DNA constructs of the present invention; plants derived therefrom having modified β-galactosidase gene expression; and seeds produced from such plants.

The β-galactosidase II protein of the present invention has demonstrated enzyme activity in cell wall disassembly leading to loss of tissue integrity and fruit softening. The β-galactosidase II protein also may be involved in cell wall turnover, which could be involved in cell extension and/or expansion and therefore plant growth and development.

By hydrolyzing galactose from the cell wall, the enzyme may allow ripening to commence and/or progress, since galactose may be involved in stimulating ethylene production alone or in conjunction with unconjugated N-glycans.

The β-galactosidase of the invention may be involved in conversion of chloroplasts (green – chlorophyll) to chromoplasts (red – lycopene) during fruit ripening by degrading chloroplast membrane galactolipids.

The family of genes represented by the nucleotide sequences shown in Figure 2 is expected to code for a group of similar enzymes with the same type of hydrolytic activity but with different tissue and/or substrate specificity's or cellular compartmentation profiles.

The β-galactosidase II protein of the present invention as well as other proteins encoded in the nucleotide sequences shown in Figure 2 may be used for preparation of pectin and other cell wall derived polymers with lowered galactosyl content for use in biofilms and solutions (for example in

clarification of fruit juices) requiring lower or higher cross-linking or

viscomertric properties.

The present invention also provides β-galactosidase enzymes for use as

components of enzyme mixtures for protoplast isolation.

5

Brief Description of the Figures

Figure 1A and 1B shows a phylogenic tree based on shared amino acid

sequence identity among tomato β-galactosidase clones TGB1-7 and other

known plant β-galactosidase polypeptides.

10

Figure 2 shows cDNA sequences [SEQ ID NOs: 1-7, respectively] for the

seven β-galactosidase genes of the invention: TGB1, TGB2, TGB3, TGB4,

TGB5, TGB6, TGB7.

15 **Figure 3** shows multiple sequence alignment of the deduced amino acid

sequences of tomato fruit for cDNA clones TGB1, TGB2, TGB3, TGB4,

TGB5, TGB6 and TGB7 [SEQ ID NOs: 8-16, respectively] and various plant

β-galactosidase cDNA clones.

20 **Figure 4** shows autoradiograph of northern blot analysis of TBG expression in

various plant tissues (flowers, leaves, roots and stems).

Figure 5 shows Autoradiograph of northern blot analysis of TBG expression

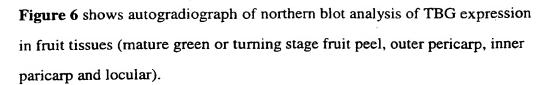in fruit tissues at different stages of development.

25

Figure 6 shows autogradiograph of northern blot analysis of TBG expression in fruit tissues (mature green or turning stage fruit peel, outer pericarp, inner paricarp and locular).

5   **Figure 7** shows autoradiograph of northern blot analysis of TBG expression in normal and mutant fruit tissues.

Figure 8 shows autoradiograph of northern blot analysis of TBG expression in response to ethylene treatment of mature green fruit tissues.

10   **Figure 9** shows Western blot analysis of TBG4 expression by yeast.

Figure 10 shows detection of β-galactosidase activity from pZBG2-1-4 expression in *E. coli*.

15   **Figure 11 A - E (1-4)** shows the comparative results of texture measurements for fruit from tomato plants containing antisense constructs to suppress TBG4 mRNA and fruit from the parental line.

20   **Figures 12A - B** show Northern blot analysis of TBG4 expression in transgenic fruit containing TBG4 antisense construct.

Figure 13 shows a Binary construct used to transform plants and express TBG4 (pZBG2-1-4) in the antisense orientation.

25

Detailed Description

The following detailed description is directed to a preferred

embodiment of the present invention and is intended as illustrative of each of

other DNA sequences of the present invention.

5   The present invention provides isolated nucleic acid molecules

comprising a polynucleotide encoding β-galactosidase polypeptides,

particularly a β-galactosidase II polypeptide having the amino acid sequence

shown in Figure 2. The DNA sequence of the exemplary β-galactosidase II

cDNA clone of the invention, which was determined from a cDNA clone,

10   pZBG2-1-4, encoding β-galactosidase II, is recorded in GenBank as Accession

Number AF020390. Not all β-galactosidases possess *in vitro* activity against

extracted cell wall material via the release of galactose from wall polymers

containing $\beta(1\rightarrow4)$-D-galactan. The polypeptide expressed from the exemplary

β-galactosidase II clone, pZBG2-1-4, has been shown to exhibit

15   β-galactosidase activity and exogalactinase activity.

The exemplary β-galactosidase II protein of the present invention, as

shown in Figure 2, shares sequence homology with the amino acid sequence

deduced from β-galactosidase cDNA clones of TBG2-7 and cDNA clones of

the fruits of asparagus (accession number P45582), apple (accession number

20   P48981), and carnation (accession number Q00662), as well as with β-

galactosidase cDNA clones of a previously published sequence of a tomato β-

galactosidase cDNA clone designated pTomβgal1 (accession number P48980)

isolated from ripe 'Ailsa Craig' fruit (Carey *et al.*, 1995). The ORF of the

clone TBG1 disclosed herein by the inventors (accession number AF023847)

is nearly identical to the cDNA previously described by Carey et al. As shown

in Figure 2, the shared deduced sequence identity is high among all the

published plant β-galactosidases of the seven clones (TBG1-7) and the other

plant β-galactosidases.

5   BLAST searches of the database also indicated significant shared

sequence identity between domains of the plant β-galactosidases and

mammalian and fungal β-galactosidases, however little share sequence identity

was detected with bacterial β-galactosidases.

As shown in Figure 1, the shared amino acid identity of TBG1 and

10   TBG3 was high. TBG4 was also very similar to both TBG1 and 3. The amino

acid sequences of TBG2 and 7 were unique because several regions of amino

acid insertions appear throughout their sequence (Figure 3).

Nucleic Acid Molecules

Unless otherwise indicated, all nucleotide sequences determined by

15   sequencing a DNA molecule herein were determined using a PCR-based

dideoxynucleotide terminator protocol and an ABI automated DNA sequencer

(such as the Model 373 from Applied Biosystems, Inc., Foster City, CA), and

all amino acid sequences of polypeptides encoded by DNA molecules

20   determined herein were predicted by translation of a DNA sequence

determined as above. Therefore, as is known in the art for any DNA sequence

determined by this automated approach, any nucleotide sequence determined

herein may contain some errors. Nucleotide sequences determined by

automation are typically at least about 90% identical, more typically at least

about 95% to at least about 99.9% identical to the actual nucleotide sequence of the sequenced DNA molecule. The actual sequence can be more precisely determined by other approaches including manual DNA sequencing methods well known in the art. As is also known in the art, a single insertion or

5 deletion in a determined nucleotide sequence compared to the actual sequence will cause a frame shift in translation of the nucleotide sequence such that the predicted amino acid sequence encoded by a determined nucleotide sequence will be completely different from the amino acid sequence actually encoded by the sequenced DNA molecule, beginning at the point of such an insertion or

10 deletion.

By "nucleotide sequence" of a nucleic acid molecule or polynucleotide is intended, for a DNA molecule or polynucleotide, a sequence of deoxyribonucleotides, and for an RNA molecule or polynucleotide, the corresponding sequence of ribonucleotides (A, G, C and U), where each

15 thymidine deoxyribonucleotide (T) in the specified deoxyribonucleotide sequence is replaced by the ribonucleotide uridine (U).

Using the information provided herein, such as the exemplary nucleotide sequence shown in Figure 2 [SEQ ID NO: 4], a nucleic acid molecule of the present invention encoding a β-galactosidase II polypeptide may be obtained

20 using standard cloning and screening procedures, such as those for cloning cDNAs using mRNA as starting material. Illustrative of the invention, the nucleic acid molecule described in Figure 2 [SEQ ID NO: 4] was discovered in a cDNA library derived from breaker, turning and pink fruit pericarp from 'Rutgers' tomato plants.

The complete sequence of the cDNA insert of pZBG2-1-4 is accessible in

the GenBank (no. AF020390) and is provided in Figure 2 [SEQ ID NO: 4].

The cDNA insert is 2532 nucleotides (nt) long and contains a single, long open

reading frame (ORF) predicted to start with the first in-frame ATG at nt 64

5   and end with TAA at nt 2238. This ORF codes for a 79 kD protein 724 amino

acids long. The deduced amino acid sequence of pZBG2-1-4 shared

significant amino acid identity to all published plant β-galactosidase sequences

in the database (Figure 1A,B). When the entire ORF of each β-galactosidase

gene was compared to pZBG2-1-4, the shared sequence identity was about

10   64% for tomato pTomβgal 1 (P48980), about 67.6% for apple (P48981), about

63% for asparagus (P45582) and about 55% for carnation (Q00662). As one

of ordinary skill would appreciate, due to the possibilities of sequencing errors

discussed above, the actual complete β-galactosidase II polypeptide encoded

by the deposited cDNA, which comprises about 724 amino acids, may be

15   somewhat longer or shorter. More generally, the actual open reading frame

may be anywhere in the range of ±20 amino acids, more likely in the range of

±10 amino acids, of that predicted from either the first methionine codon from

the N-terminus shown in Figure 2 [SEQ ID NO: 4]. In any event, as discussed

further below, the invention further provides polypeptides having various

20   residues deleted from the N-terminus of the complete polypeptide, including

polypeptides lacking one or more amino acids from the N-terminus of the β-

galactosidase II polypeptide described herein.

Leader and Mature Sequences

Analysis of the deduced amino acid sequence of pZBG2-1-4 suggested a high probability for secretion based on the presence of a hydrophobic leader sequence, a leader sequence cleavage site and three possible N-glycosylation

5   sites. The programs PSORT V6.4 (Nakai and Kanehisa, 1992, incorporated herein by reference) and SignalP V1.1 (Nielsen et al., 1997, incorporated herein by reference), were used to predict that the ORF contains a hydrophobic leader sequence that would be cleaved between the alanine and serine residues at positions 23 and 24 respectively, and that the mature polypeptide has an

10   extracellular location. The mature polypeptide contains three possible N-glycosylation sites at asparagine numbers 282, 459 and 713, however the asparagine at position 713 is unlikely to be glycosylated due to the proline at position 714. The predicted molecular mass of the unglycosylated mature polypeptide was 75 kD with a pI of 8.9.

15   Accordingly, the amino acid sequence of the complete β-galactosidase II protein of the invention includes a leader sequence and a mature protein, as shown in Figure 3 [SEQ ID NO: 4]. More in particular, the present invention provides nucleic acid molecules encoding a mature form of the β-galactosidase II protein. Thus, according to the signal hypothesis, secreted proteins have a

20   signal or secretory leader sequence which is cleaved from the complete polypeptide to produce a secreted "mature" form of the protein. In some cases, cleavage of a secreted protein is not entirely uniform, which results in two or more mature species of the protein. Further, it has long been known that the cleavage specificity of a secreted protein is ultimately determined by the

primary structure of the complete protein, that is, it is inherent in the amino

acid sequence of the polypeptide. Therefore, the present invention provides a

nucleotide sequence encoding the mature β-galactosidase II polypeptide

having the amino acid sequence encoded by the cDNA shown in Figure 2

[SEQ ID NO: 4] and provided in GenBank (Accession No. AF20390). By the

"mature β-galactosidase II polypeptide having the amino acid sequence

encoded by the cDNA clone shown in Figure 2 [SEQ ID NO: 4] is meant the

mature form(s) of the β-galactosidase II protein produced by expression in a

plant cell of the complete open reading frame encoded by the cDNA sequence

of the clone shown in Figure 2 [SEQ ID NO: 4] and provided in GenBank

(Accession No. AF20390).

The exemplary β-galactosidase II cDNA of the present invention

(TBG4) has been expressed in *E. coli* strain XLI blue MR (lacZ) (Stratagene,

La Jolla, CA), as described hereinbelow (see Example).

Analysis of the deduced amino acid sequence of cDNA clones

representing the other β-galactosidase genes of the invention also revealed

open reading frames and, in some cases, suggested a high probability for

secretion of the encoded proteins. All the full-length cDNA clones were

predicted to have a signal sequence (Fig. 2). Using the two prediction

programs SignalP and PSORT, TBG4 was predicted to be secreted by both

programs. TBG1, 2 and 3 were predicted to have cleavable signal sequences

by SignalP, but uncleavable signal sequences by PSORT. TBG7 was suggested

to be targeted to the chloroplast by PSORT. Particular observations for each

of the seven clones are as follows, based on the presence of a hydrophobic

leader predicted by the programs PSORT V6. and SignalP V1.1: TBG1:

initiation codon at 306 [SEQ ID NO: 1], ORF = 835 amino acids [SEQ ID

NO: 8], signal sequence at 1-24; TBG2: initiation codon not determined [SEQ

ID NO: 2], ORF = 888 amino acids [SEQ ID NO: 9], signal sequence at 1-25;

5   TBG3: initiation codon at 32 [SEQ ID NO: 3], ORF = 838 amino acids [SEQ

ID NO: 10], signal sequence at 1-22; TBG5: initiation codon not determined

[SEQ ID NO:5], ORF = 251 amino acids [SEQ ID NO: 12], signal sequence

not determined; TBG6: initiation codon not determined [SEQ ID NO:6], ORF

= 248 amino acids [SEQ ID NO:13], signal sequence not determined; TBG7:

10   initiation codon at 104 [SEQ ID NO: 7], ORF = 870 amino acids [SEQ ID

NO:14], signal sequence at 1-35.

The deduced amino acid sequences of the seven clones was also

subjected to analysis using the program DNAsis and the predictions for

molecular mass, cellular targeting, pI and potential N-linked glycosylation

15   sites are summarized in Table I.

Table I. Tomato β-galactosidase (TBG) cDNA sequence data. **Fiv full-length and 2 partial-length cDNAs were cloned and sequenced. The DNA and deduced amino acid sequence data is presented below**

CLONE	mRNA(kb)	kD	pl	N-LINK	TARGET
TBG1	3.2	90.8	6.2	2	ER/OUT
TBG2	3.0	97.0	6.2	6	PM
TBG3	2.8	90.5	8.2	1	ER/OUT
TBG4	2.6	77.9	8.9	3	OUT
TBG5	~3				
TBG6	~3				
TBG7	3.0	93.3	8.0	6	CHLOR

N-LINK = possible N-linked glycosylation sites; ER = endoplasmic reticulum; out = secreted; PM = tethered to plasma membrane; CHLOR = chloroplast

As indicated, nucleic acid molecules of the present invention may be in the form of RNA, such as mRNA, or in the form of DNA, including, for instance, cDNA and genomic DNA obtained by cloning or produced synthetically. The DNA may be double-stranded or single-stranded. Single-stranded DNA or RNA may be the coding strand, also known as the sense strand, or it may be the non-coding strand, also referred to as the anti-sense strand.

By "isolated" nucleic acid molecule(s) is intended a nucleic acid molecule, DNA or RNA, which has been removed from its native environment

For example, recombinant DNA molecules contained in a vector are

considered isolated for the purposes of the present invention. Further

examples of isolated DNA molecules include recombinant DNA molecules

maintained in heterologous host cells or purified (partially or substantially)

5   DNA molecules in solution. Isolated RNA molecules include *in vivo* or *in*

vitro RNA transcripts of the DNA molecules of the present invention. Isolated

nucleic acid molecules according to the present invention further include such

molecules produced synthetically.

Isolated nucleic acid molecules of the present invention include DNA

10   molecules comprising an open reading frame (ORF) with an initiation codon at

position 64 of the nucleotide sequence shown in Figure 2 [SEQ ID NO: 4].

Also included are DNA molecules comprising the coding sequence for the

mature β-galactosidase II protein shown at positions 135-2532 of Figure 2

[SEQ ID NO: 4].

15   In addition, isolated nucleic acid molecules of the invention include

DNA molecules which comprise a sequence substantially different from those

described above but which, due to the degeneracy of the genetic code, still

encode the β-galactosidase II protein. Of course, the genetic code and species-

specific codon preferences are well known in the art. Thus, it would be

20   routine for one skilled in the art to generate the degenerate variants described

above, for instance, to optimize codon expression for a particular host (e.g.,

change codons in the plant mRNA to those preferred by a bacterial host such

as *E. coli*). Preferably, this nucleic acid molecule will encode the mature

polypeptide encoded by the above-described deposited cDNA clone.

The invention further provides an isolated nucleic acid molecule

having the nucleotide sequence shown in Figure 2 [SEQ ID NO: 4] or a

nucleic acid molecule having a sequence complementary to the above

sequence. Such isolated molecules, particularly DNA molecules, are useful as

5   probes for gene mapping, by in situ hybridization with chromosomes, and for

detecting expression of the β-galactosidase II gene in plant tissue, for instance,

by Northern blot analysis.

The present invention is further directed to nucleic acid molecules

encoding portions of the nucleotide sequences described herein as well as to

10   fragments of the isolated nucleic acid molecules described herein. In

particular, the invention provides a polynucleotide having a nucleotide

sequence representing the portion of Figure 2 [SEQ ID NO: 4] which consists

of positions 1-2538 of Figure 2 [SEQ ID NO: 4].

In addition, the invention provides additional nucleic acid molecules

15   having nucleotide sequences related to extensive portions of Figure 2 [SEQ

ID NO: 4] which have been determined from the following related cDNA

clones: TBG1-3 and TBG5-7 as shown in Figure 3, SEQ. NO's 1-3 and 5-7

In another aspect, the invention provides an isolated nucleic acid

molecule comprising a polynucleotide which hybridizes under stringent

20   hybridization conditions to a portion of the polynucleotide in a nucleic acid

molecule of the invention described above, for instance, the cDNA clone

shown in Figure 2 [SEQ ID NO: 4]. By "stringent hybridization conditions" is

intended overnight incubation at 42° C in a solution comprising: 50%

formamide, 5x SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium

25   phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 μg/ml

denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65° C.

As indicated, nucleic acid molecules of the present invention which encode a β-galactosidase II polypeptide may include, but are not limited to those encoding the amino acid sequence of the mature polypeptide, by itself; and the coding sequence for the mature polypeptide and additional sequences, such as those encoding the about 1-23 amino acid leader sequence, such as a pre-, or pro- or prepro- protein sequence; the coding sequence of the mature polypeptide, with or without the aforementioned additional coding sequences.

Also discovered is an enhancer/promoter associated with expression of the genes encoding β-galactosidase. The inventors have characterized the expression profile of TBG2 mRNA and have cloned a lambda genomic cDNA. TBG2 is expressed before the onset of fruit ripening and continues at uniform level throught all the ripening stages. TBG2 has been found to be expressed in all fruit tissues and has also been found to be fruit specific. Experiments have shown TBG2 to be unaffected by ethylene. TBG2 is expressed in the ripening mutants rin, nor and Nr at the normal chronological time after anthesis. The promoter discovered would be useful to express any gene in the sense or antisense orientation, specifically in tomato fruit, in all tomato fruit tissues, starting before and continuing throughout the entire ripening process. The promoter could also be used to express any gene in the ripening mutants rin, nor and Nr without the need to gas the fruit with exogenous ethylene.

Variant and Mutant Polynucleotides

The present invention further relates to variants of the nucleic acid

molecules of the present invention, which encode portions, analogs or

derivatives of the β-galactosidase II protein. Variants may occur naturally,

such as a natural allelic variant. By an "allelic variant" is intended one of

several alternate forms of a gene occupying a given locus on a chromosome of

an organism. *Genes II*, Lewin, B., ed., John Wiley & Sons, New York (1985).

Non-naturally occurring variants may be produced using art-known

mutagenesis techniques.

Such variants include those produced by nucleotide substitutions,

deletions or additions. The substitutions, deletions or additions may involve

one or more nucleotides. The variants may be altered in coding regions,

non-coding regions, or both. Alterations in the coding regions may produce

conservative or non-conservative amino acid substitutions, deletions or

additions. Especially preferred among these are silent substitutions, additions

and deletions, which do not alter the properties and activities of the β-

galactosidase II protein or portions thereof. Also especially preferred in this

regard are conservative substitutions.

Most highly preferred are nucleic acid molecules encoding the mature

protein having the amino acid sequence shown in Figure 2 as pZBG2-1-4 or

the mature β-galactosidase II amino acid sequence encoded by the deposited

cDNA clone.

Further embodiments include an isolated nucleic acid molecule

comprising a polynucleotide having a nucleotide sequence at least 90%

identical, and more preferably at least 95%, 96%, 97%, 98% or 99% identical

to a polynucleotide selected from the group consisting of: (a) a nucleotide

sequence encoding the β-galactosidase II polypeptide having the complete

amino acid sequence in Figure 2 [SEQ ID NO: 4] (b) a nucleotide sequence

5   encoding the mature β-galactosidase II polypeptide shown in Figure 2 [SEQ

ID NO: 4]; (c) a nucleotide sequence complementary to any of the nucleotide

sequences in (a) or (b) above.

Vectors and Host Cells

The present invention also relates to vectors which include the isolated

10   DNA molecules of the present invention, host cells which are genetically

engineered with the recombinant vectors, and the production of β-

galactosidase II polypeptides or fragments thereof by recombinant techniques.

The vector may be, for example, a phage, plasmid, viral or retroviral vector.

Retroviral vectors may be replication competent or replication defective. In

15   the latter case, viral propagation generally will occur only in complementing

host cells.

The polynucleotides may be joined to a vector containing a selectable

marker for propagation in a host. Generally, a plasmid vector is introduced in

a precipitate, such as a calcium phosphate precipitate, or in a complex with a

20   charged lipid. If the vector is a virus, it may be packaged in vitro using an

appropriate packaging cell line and then transduced into host cells.

The DNA insert should be operatively linked to an appropriate

promoter, such as the phage lambda PL promoter, the *E. coli lac, trp, phoA*

and *tac* promoters, the SV40 early and late promoters and promoters of

retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a translation initiating codon at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase, G418 or neomycin resistance for eukaryotic cell culture and tetracycline, kanamycin or ampicillin resistance genes for culturing in *E. coli* and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as *E. coli*, StrepZBG2-1-4yces and *Salmonella typhimurium* cells; fungal cells, such as yeast cells; insect cells such as Drosophila S2 and Spodoptera Sf9 cells; animal cells such as CHO, COS, 293 and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from QIAGEN, Inc., *supra*; pBS vectors, Phagescript vectors, Bluescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

Introduction of the construct into the host cell can be effected by

calcium phosphate transfection, DEAE-dextran mediated transfection, cationic

lipid-mediated transfection, electroporation, transduction, infection or other

methods. Such methods are described in many standard laboratory manuals,

5   such as Davis *et al.*, *Basic Methods In Molecular Biology* (1986).

Example

Tomato (*Lycopersicon esculentum* Mill., cv. 'Rutgers') plants were grown

in a greenhouse using standard cultural practices. The ripening mutants,

10   *ripening inhibitor* (*rin*), *non-ripening* (*nor*) and *never ripe* (*Nr*) (Tigchelaar *et*

al., 1978), were all in the 'Rutgers' background. Flowers were tagged at

anthesis and fruit were harvested according to the number of days post-

anthesis (dpa) or based on their surface color using ripeness stages as

previously described (Mitcham *et al.*, 1989), the complete disclosure of which

15   is hereby fully incorporated herein by reference. For gene expression studies,

a variety of leaf, flower, and stem tissues were harvested from greenhouse-

grown plants and roots were harvested from seedlings grown in basal tissue

culture medium for 4 weeks after seed germination.

20   **RNA Extraction**

Fruits were processed immediately after harvest in the greenhouse by

chilling on ice, excising the various tissues and freezing them in liquid

nitrogen. Tissue samples were ground using a mortar and pestle and stored at -

80°C. RNA was extracted using the method described in Verwoerd et al.

25   (1989). Poly(A)RNA was purified from total RNA using oligo(dT) columns

(Pharmacia, Piscataway, NJ). RNA was quantified by measuring A_{260} using a

dual beam spectrophotometer.

RT-PCR

Degenerate primers were designed based on the highest shared deduced

amino acid sequence identity we found between an apple (accession number

P48980), asparagus (P45582) and carnation (Q00662) β-galactosidase cDNA

clones. The two primers used for the first reaction were BG5'E1

(WSNGGNWSNATHCAYTAYCC) and BG3'E

(CCRTAYTCRTCNADNGGNGG). A second reaction was done on the

products of the first reaction using BG5'I1

(ATHCARACNTAYGTNTTYTGG) and BG3'E. The degeneracy code for the

primer sequences is N=a+t+c+g; H=a+t+c; B=t+c+g; D=a+t+g; V=a+c+g;

R=a+g; Y=c+t; M=a+c; K=t+g; S=c+g; and W=a+t. The 5' and 3' primers

corresponded to amino acids 72-78 and 321-315 of the apple clone,

respectively. Amplification was done using AmpliTaq DNA polymerase

(Perkin Elmer, Norwalk, CT) and standard PCR conditions using the cDNA

made for the first cDNA library described below as a template (Ausubel *et al.*,

1987). PCR products were separated in an agarose gel and fragments of the

expected size (approximately 750 bp) were purified, cloned into pCRscript

(Stratagene, La Jolla, CA) , and sequenced.

cDNA library

Two cDNA libraries were constructed. The first comprised poly(A) RNA

isolated from breaker, turning and pink fruit pericarp from 'Rutgers' plants.

The cDNA synthesis and library construction was done exactly according to

the manufacturers instructions for the ZAP-cDNA Gigapack II Gold Cloning

Kit (Stratagene), the complete disclosure of which is fully incorporated herein

by reference. First-strand cDNA synthesis was primed using a poly(dT)

5   primer and inserts were directionally cloned into the Uni-Zap XR vector using

EcoRI and XhoI restriction sites. The second library comprised poly(A) RNA

isolated from all fruit tissues (except seeds) from immature green, mature

green, breaker, turning, pink, red-ripe and over-ripe fruit of 'Rutgers' plants.

The cDNA synthesis and library construction was done exactly according to

10   the manufacturers instructions for the SuperScript Lambda System for cDNA

synthesis and • Cloning (GibcoBRL, Gaithersburg, MD). First-strand cDNA

synthesis was primed using a oligo(dT) primer and cDNA inserts were

directionally cloned into the • ZipLox cloning vector using SalI and NotI

restriction sites. Both libraries were amplified and maintained using the host

15   strains provided by the manufacturer, according to their instructions.

One of the clones (RT-PCR2-1) was used to screen 10^6 plaques from the

tomato fruit cDNA libraries at low stringency (hybridization at 45°C, no

formamide and final wash with 0.2X SSC at 42°C). Thirty positive cDNA

clones were identified and partially sequenced. Complete sequencing and

20   characterization of the RT-PCR and cDNA clones revealed the possibility of

seven unique β-galactosidase genes.

DNA and RNA Gel Blot Analysis

Southern analysis was done using the 3' UTR of each full length clone

and the RT-PCR clones as probes against restriction enzyme digested genomic

DNA. DNA gel blot analysis was done essentially as described in Smith and

5   Fedoroff (1995) except that 3 µg of genomic DNA was used for each digest.

The genes corresponding to the clones appeared to be present as single copies

(data not shown). The same probes were used to map 6 of the 7 genes using

RFLPs of recombinant inbred lines and the loci names and map positions are

shown in Table II (James Gioviannone, Texas A&M University, personal

10   communication).

Table II. TBG loci map positions. Genes were maped by Southern
analysis using RFLPs of recombinant inbred lines.

Gene	chromosome	map position
TBG1	12*	overlap of IL 12-2, IL 12-3
TBG2	9	IL 9-3
TBG3	3	IL 3-5
TBG4	12*	overlap of IL 12-2, IL 12-3
TBG5	11	IL 11-3
TBG6	2	overlap of IL 2-4, IL 2-5
TBG7	no RFLP	

*TBG1 and 4 are loosely linked

Total RNA (20 µg/ lane) was separated in a formaldehyde/Mops

agarose gel, transferred to Hybond-N$^+$ nylon membrane (Amersham, Arlington

15   Heights, IL), fixed by incubating for 2 h at 80°C, hybridized overnight in a

hybridization incubator (Robbins Scientific, Sunnyvale, CA) using a buffer

described by Church and Gilbert (1984) washed to a final stringency of 0.1 X

SSC with 0.2% SDS at 65°C, and autoradiographed essentially as described by

Ausubel et al. (1987). An RNA ladder standard (GibcoBRL) was used to

estimate the length of the RNAs. Probes were synthesized using a random

priming kit with ^{32}P-dATP as the label (Boehringer Mannheim, Indianapolis,

IN). Northern analysis was done using the 3' UTR of each full length clone

and the RT-PCR clones as templates for probe synthesis. As a loading control,

RNA blots were stripped and re-probed at a reduced hybridization and

washing stringency using a soybean 26S rDNA fragment (Turano et al., 1997).

For all hybridizations, ^{32}P(dATP)-labeled probe was diluted to 1-2 x 10^6

dpm/mL. The complete disclosures of the above references are fully

incorporated herein by reference.

Sequence Analysis

Sequencing was done at the Iowa State University Sequencing Facility

(Ames, IA) using a PCR-based dideoxynucleotide terminator protocol and an

ABI automated sequencer (Applied Biosystems, Foster City, CA). The

sequencing of both cDNA insert strands was done by primer walking.

Nucleotide and deduced amino acid sequence comparisons against the

databases were done using BLAST searches (Altschul et al., 1990). Sequence

data were analyzed and aligned using DNA Strider 1.2 (Marck, 1988) and

MacDNAsis (Hitachi, San Bruno, CA) software. The complete disclosures of

the above references are fully incorporated herein by reference.

Northern Blot Analysis

Tissue Specific Expression

Northern blot analysis was done to reveal which, if any, of the β-galactosidase genes had a fruit-specific expression pattern. With the exception

5   of TBG2, transcripts of all clones were detected in non-fruit tissues (Fig. 4).

Transcripts of TBG 1, 4, 5 and 6 were detected in all the tissues tested. TBG3

transcript was detected at low levels in root and stem tissues, while TBG7

transcript was detected in flower and stem tissues.

Temporal expression pattern in fruit

10

The temporal expression pattern of the seven genes in fruit tissue was

examined using RNA extracted from all fruit tissues except seeds. Transcripts

for all seven genes were detected during some stage of fruit development (Fig.

5). TBG1 and 3 had similar expression patterns and their transcripts were

15   detected throughout the breaker to over-ripe stages. TBG2 had a unique

expression pattern and its transcript was detected at a constant level from 30

dpp to the over ripe stage. TBG4 expression pattern was similar to TBG1 and

3, but differed in that the transcript level was significantly higher at the turning

stage. TBG5 had a similar expression pattern to TBG4 during the ripening

20   stages of development, however TBG5 transcript was also detected throughout

all the earlier stages of fruit development. TBG6 had an interesting expression

pattern and its transcript was only detected at high levels in all pre-ripening

stages tested. TBG7 also had a unique expression pattern and its transcript was

detected at very low levels throughout all the stages tested, and at moderate

25   levels at 10 dpp and the over-ripe stage.

Spatial expression pattern in fruit

Northern blot analysis was also done to determine transcript accumulation in various fruit tissues. Since there were temporal differences in the expression patterns of the TBG genes both the mature green and turning fruit stages were used for RNA extractions (Fig. 6). Both TBG2 and TBG6 transcripts were detected in all mature green fruit tissues tested. TBG7 transcript was present in all fruit tissues tested except for locules. Both TBG1 and TBG4 transcripts were detected in RNA samples extracted from all turning stage fruit tissues. TBG4 transcript was notably more abundant in the peel. TBG3 and TBG5 expression patterns were unique and their transcripts were detected in all tissues except the outer pericarp and locular respectively.

Expression in normal versus mutant fruit

In order to better understand the potential roles of the TBG products and transcriptional regulatory mechanisms, northern analysis was performed using fruit tissue from the ripening mutants *rin, nor* and N^r. This analysis was important because it might give clues for preliminary determination of any potential ripening and/or softening role any of the TBGs might possess.

The results of mutant fruit Northern analysis suggested that the transcriptional regulation of TBG1, 2, 3, 5 and 7 was unaffected in mutant fruit tissue and that their transcripts were present in a normal chronological (dpp) pattern (Fig. 7). The abundance of TBG4 and 6 transcripts were however different in the mutant fruit. TBG4 transcript was not detected in fruit tissue of N^r and was detected at much lower levels in *rin* and *nor* than wild type fruit

tissues. Normally TBG6 transcripts are detectable at high levels throughout the

early stages of fruit development but are not detectable after the mature green

stage (40-42 dpp). TBG6 transcripts persisted even to 50 dpp in fruit of all

three mutants.

5

Transcriptional regulation by ethylene

The northern analysis done using mutant and wild type fruit suggested

that TBG4 expression might be up-regulated by ethylene and that TBG6

expression might be down-regulated by ethylene. In order to evaluate this

10   hypothesis mature green fruit were harvested and subjected to a continuous

flow of 10 ppm ethylene mixed in air. Control and ethylene-treated fruit were

used for RNA extractions at 1, 2, 12 and 24 hours. The results of this

experiment confirmed the findings from the mutant fruit northern analysis. As

expected, the presence and abundance of TBG1, 2, 3, 5 and 7 transcripts was

15   essentially unaffected in mature green tissues subjected to exogenous ethylene

treatment (Fig. 8). However, TBG4 transcript abundance was increased in

mature green tissues in the presence of ethylene. From the data presented it

was unclear whether TBG6 transcript abundance was reduced by exogenous

ethylene treatment since its transcript level was normally reduced at this stage

20   of fruit development.

Enzyme activity

In order to determine the role of the TBG encoded products we

initiated experiments to express the cDNA encoded enzymes using

25   heterologous expression systems. Several E. coli expression systems were

tested, but the yield of product was very low due to toxicity (See the example

below). Therefore we used a yeast expression system which secretes a mature

amino-terminal-FLAG fusion protein into the culture medium. The TBG4

cDNA was tested first and resulted in the production of approximately 1 mg

5   TBG4 active protein per 50 mls culture. TBG4 was used first because the

cDNA codes for the enzyme β-galactosidase II which was purified from

tomato fruit and has been characterized in some detail (Carey et al 1995, Smith

et al 1998). Therefore we could compare the activity of the heterologous

system-expressed protein to the native enzyme purified from tomato. The

10   TBG4 protein was successfully affinity purified using an anti-FLAG affinity

resin (Figure 9).

The affinity-purified TBG4 enzyme was shown to have β(1→4)-D-

galactosidase activity by virtue of its ability to hydrolyze the synthetic

substrate p-nitrophenyl-β-D-galactopyranoside (Smith et al. 1998). The

15   enzyme can cleave galactosyl residues from a variety of cell wall substrates

and therefore has exo-galactanase activity (Table III). The remaining full-

length cDNA clones are currently being tested for successful expression of

active enzyme. Preliminary results have shown that TBG1 codes for an

enzyme which also has both β-D-galactosidase and exo-galactanase activity

20   (Table III).

Table III. Cell wall degrading activity of TBG4 and TBG1 expressed in yeast. Removal of galactosyl residues from chelator soluble (CSP) and alkali soluble (ASP) pectin and hemicellulosic (HCF) cell wall fractions purified from tomato fruit.

enzyme	substrate	μg galactose released	
		boiled	live
[a]TBG4	CSP	0	5
	ASP	0	14.5
	HCF	0	4
[b]TBG1	ASP	0	1.2

2 mg substrate; 4 hours at 37°C
[a].005 units enzyme/rx
[b].0005 units enzyme/rx

pZBG2-1-4 Codes for a β-Galactosidase

The TBG4 ORF was cloned in-frame into the repressible/inducible

bacterial expression vector pFLAG-CTC. The host strain XL1-Blue MR is a

mutant strain containing no endogenous β-galactosidase activity nor α-

complementation. Induction of gene transcription by (IPTG) caused the

immediate cessation of *E. coli* growth at 30 to 37°C. However, induction at

20°C did allow for some limited *E. coli* growth. When clones containing the

pZBG2-1-4 4 ORF were grown at 20°C and induced with IPTG, the cells

slowly turned blue after 36 hrs growth in medium containing the β-

galactosidase substrate X-GAL, (Figure 10). If not induced with IPTG, no

blue color was seen, even after extended growth in media containing X-GAL.

As an additional negative control, clones consisting of XL1-Blue MR

transformed with the FLAG vector alone never showed any β-galactosidase

activity with or without IPTG-induction, even after 7-days growth (Fig 10).

As a positive control for maximal β-galactosidase (derived from *E. coli* β-

galactosidase) activity the cloning vector pGEM was transformed into the host

strain DH5α and the results are also shown in Figure 10. Figure 10 shows the

detection of β-galactosidase activity from pZBG2-1-4 expression in *E. coli*.

5   Cells were harvested and extracts were prepared every 12 hours and the A_{615}

measured. Cultures were grown with the addition of the chromogenic substrate

X-GAL (open symbols) or X-GAL and the transcriptional inducer IPTG

(closed symbols) in the medium. The vector used as a positive control for *E.*

coli β-galactosidase activity was pGEM (■) and the vector used as a negative

10   control and for expression was pFLAG-CTC either without (○,●) or containing

the pZBG2-1-4 ORF (△,▲).

Effects on Plant Tissue Texture

To further demonstrate the function of TBG4 encoded β-galactosidase II

15   the following experiments were carried out.

Fruit from tomato plants containing antisense constructs to suppress

TBG4 mRNA were up to 40% firmer [compare means of parental line #1 with

antisense line #2 in Figures 11A – 11E(1-4)] than fruit from the parental line.

Among the transformants the line with the firmest fruit also had the lowest

20   overall levels of TBG4 mRNA (Figure 12A,B). This correlation suggests that a

reduction in TBG4 mRNA is associated with increased fruit firmness. Firmer

fruit might result in (1) less shipping damage (a) less loss due to damage and

(b) ability to harvest at later stage resulting in better flavor at market (2) longer

shelf life for both market and consumer. (3) better quality fruit for fresh slice

market; fruit cut better at the pink/red stage when firmer.

Methods

5   To determine the function of TBG4 encoded β-galactosidase II, antisense

constructs were made using the constitutively expressed 35S CaMV promoter

to express TBG4 antisense RNA (Figure 13). Constructs were moved into

tomato using Agrobacterium-mediated transformation. Four tomato cultivars

have been transformed in order to evaluate the effect of TBG4 suppression on

10   processing tomato (cv 'UC82b') fruit paste quality and three fresh pick

cultivars. Of the fresh pick cultivars one is a soft fruit large cherry tomato (cv

'Ailsa Craig'), the second is a soft fruit old breeding line (cv 'Rutgers') and

the third is a recently developed somewhat firm cultivar 'New Rutgers'.

Among the lines where TBG4 mRNA is suppressed we expect to observe an

15   increase in firmness and paste viscosity.

Texture

       Although this project is nearly finished the complete biochemical and

molecular analysis is not finished. The preliminary results on the analysis of

20   the 'New Rutgers' cultivar is presented in Figures 11A – E(1-4) and 12A,B. In

this example a fresh pick cultivar called 'New Rutgers' was used. Plants of the

purchased seed were grown and allowed to self and the resulting seed was

used as the parental control (line 1). Seven independent transformed plants

(lines 2-8) containing TBG4 antisense constructs were grown and allowed to

25   self. Transformation (T-DNA insertion) was confirmed by southern analysis

(data not shown). From each transformed line, five plants were grown along with 10 parental line plants. Fruit were tagged at the breaker stage (1st onset of color change) and were harvested at breaker plus 7 days. Data were taken using 15-20 fruit from each line. Each type of texture measurement was done

5   twice for each fruit and fruit were subjected to 4 types of texture measurements using a Stable Micro System's TA-XT2i texture analyzer. The 4 measurements were; 1, 2-inch flat plate compression to 3 mm (Figure 1A), 2, 4 mm spherical indenter compression to 3 mm (Figure 11B), 3, 4 mm cylindrical indenter compression to 3 mm (Figure 11C) and 4, 4 mm cylindrical indenter

10   puncture to 10 mm (Figure 11D). The summary of this data is shown in Figure 11E(1-4). In Figures 11A –E (1-4) line 1 was the parental line and lines 2-8 each represent an independent transformant containing one T-DNA copy of the TBG4 antisense construct. Statistical analysis (Duncans and Scheffé) of the data revealed that fruit from the transformed lines 3, 7 and 8 were not

15   significantly different from the parental line but that transformed lines 2, 4, 5 and 6 were significantly firmer than the parental fruit. Most noteworthy is that fruit from transformed line 2 had fruit with a mean firmness that was 40% firmer than that of the parental line (Figures 11A-D).

20   **Northern Blot Analysis**

We are currently investigating any changes in the biochemical composition of fruit where TBG4 mRNA levels have been suppressed. These experiments are designed to show a link between increased fruit firmness and

25   TBG4 mRNA suppression, TBG4 encoded enzyme activity suppression,

WO 99/64564

possible cell wall modification (e.g. increased galactosyl residue content) and a

decrease in free galactose levels during fruit ripening.

These experiments are not complete, however some preliminary

Northern blot experiments were done and the data is shown in Figure 12A,B.

5   There is no parental or azygous control fruit RNA shown in Figure 12A,B

because these plants were the last to grow,and RNA extractions are just being

done now. As a comparison of normal fruit TBG4 mRNA levels refer to

Figure 5 above. The data from Figure 5 showed that TBG4 mRNA levels are

low at the mature green stage, peak at the turning stage and are reduced at the

10   red stage. All the lines except for 2 and 3 expressed antisense TBG4 mRNA

(Figure 12A,B). The antisense transcripts appear as two bands, smaller in

length than the endogenous mRNA. The two bands probably resulted from 1,

the expected transcriptional stop signal provided by the NOS-terminator and 2,

a cryptic transcriptional stop signal in the antisense TBG4 cDNA. The most

15   notable result was in line 2 where no TBG4 mRNA was detected at the turning

stage. Line 2 also had the firmest red fruit (see Figure 11A -D). The absence of

detectable TBG4 mRNA probably was the result of cosupression of both the

endogenous and antisense mRNAs. When compared to earlier blots (e.g.

Figure 4), all of the lines appeared to have an overall reduced level of TBG4

20   mRNA, but it is impossible to assign numbers to this statement without the

parental and azygous control RNA on the same Northern blot.

The specification discloses that β-galactosidase II polypeptide is involved

in the degradation of cell wall pectin during fruit ripening. In the present

invention, the role of β-galactosidases in tomato during fruit ripening and

25   softening and the description of the cloning of a β-galactosidase cDNA clone

that codes for a $\beta(1\rightarrow4)$galactan degrading enzyme, which is expressed in ripening tomato fruit tissues, has been shown.

The present work indicates that pZBG2-1-4 is a cDNA derived from the transcript of the TBG4 gene which codes for β-galactosidase II for the

5 following reasons:

First, the deduced amino acid sequence of the highly conserved amino-terminal portion of the expected mature pZBG2-1-4 translation product matches almost exactly (28 of 30 amino acids) with the amino-terminal sequence of β-galactosidase II as purified by Carey et al. (1995) and

10 designated TOMAA. Importantly, the two amino acids (KY) in the β-galactosidase II sequence (TOMAA), that do not match the pZBG2-1-4 deduced amino acid sequence of the present invention are believed to be incorrect since all plant β-galactosidase sequences in the database and four additional β-galactosidase-related cDNAs that were identified from tomato all

15 match or have conserved substitutions with the deduced amino acid sequence of pZBG2-1-4 at these same two amino acid (ST) positions (Figure 3).

Second, the transcript detected by pZBG2-1-4 is present in normal ripening fruit at the same time that β-galactosidase II activity was detected (Figure 5; Carey et al., 1995). Moreover, little or no transcript was detected

20 in fruit at 45 and 50 dpa from the mutants nor, rin and Nr (Figure 7). This observation also coincides with the data presented by Carey et al. (1995) that β-galactosidase II activity remained at levels equal to mature green fruit and did not rise in fruit 45-65 dpa from nor or rin plants. Interestingly, Carrington and Pressey (1996) have reported that β-galactosidase II activity was only

detected in 'Rutgers' fruit after the turning stage of ripeness. The Northern

data in the present invention indicates that maximum β-galactosidase II

activity occurs only after the turning stage, assuming mRNA levels predict

extractable enzyme activity (Figure 5).

5   Third, the apparent molecular weight of 77.9 kD and pI of 8.9 for the

mature protein predicted from the pZBG2-1-4 sequence is similar to that

determined for β-galactosidase II., Pressey (1983), estimated a molecular

weight of 62 kD by gel-filtration column chromatography and a pI of 7.8 by

isoelectric focusing, while Carey *et al.* (1995) estimated a molecular weight of

10   75 kD by SDS-PAGE and a pI of 9.8 by isoelectric focusing.

Fourth, enzyme produced from pZBG2-1-4 ORF using a heterologous

yeast expression system has both β-galactosidase activity and exogalactinase

activity.

LITERATURE CITED

Altschul SF, Gish W, Miller W, Meyers EW, Lipman DJ (1990) Basic local
alignment search tool. J Mol Biol **215**:403-410

Ausubel F, Brent R, Kingston R, Moore D, Seidman J, Smith J, Struhl K,
eds, (1987) Current Protocols in Molecular Biology. John Wiley and
Sons, New York

Carey AT, Holt K, Picard S, Wilde R, Tucker GA, Bird CR, Schuch W,
Seymour GB (1995) Tomato exo-(1→4)- β-D-galactanase. Isolation,
changes during ripening in normal and mutant tomato fruit, and
characterization of a related clone. Plant Physiol **108**:1099-1107

Carrington CM, Pressey R (1996) β-galactosidase II activity in relation to
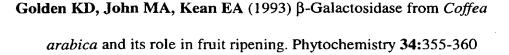changes in cell wall galactosyl composition during tomato ripening. J
Amer Soc Hort Sci **121**:132-136

Church GM, Gilbert W (1984) Genomic sequencing. Proc Natl Acad Sci
USA **81**:1991-1995

DeVeau EJ, Gross KC, Huber DJ, Watada AE (1993) Degradation and
solubilization of pectin by β-galactosidases purified from avocado
mesocarp. Physiol Plant **87**:279-285

Fischer RL, Bennett AB (1991) Role of cell wall hydrolases in fruit ripening.
Annu Rev Plant Physiol Plant Mol Bio **42**:675-703

Giovannoni JJ, DellaPenna D, Bennett AB, Fischer RL (1989) Expression
of a chimeric polygalacturonase gene in transgenic *rin* (ripening inhibitor)
tomato fruit results in polyuronide degradation but not fruit softening.
Plant Cell **1**:53-63

Golden KD, John MA, Kean EA (1993) β-Galactosidase from *Coffea arabica* and its role in fruit ripening. Phytochemistry **34**:355-360

Gross KC (1984) Fractionation and partial characterization of cell walls from normal and non-ripening mutant tomato fruit. Physiol Plant **62**:25-32

5   **Gross KC, Sams CE** (1984) Changes in cell wall neutral sugar composition during fruit ripening: A species survey. Phytochemistry **23**:2457-2461

Gross KC, Wallner SJ (1979) Degradation of cell wall polysaccharides during tomato fruit ripening. Plant Physiol **63**:117-121

Hall LN, Tucker GA, Smith CJS, Watson CF, Seymour GB, Bundick Y,

10   **Boniwell JM, Fletcher JD, Ray JA, Schuch W, Bird CR, Grierson D.** (1993) Antisense inhibition of pectin esterase gene expression in transgenic tomatoes. Plant J **3**:121-129

Huber DJ (1983) The role of cell wall hydrolases in fruit softening. Hort Rev **5**:169-219

15   **Kang IK, Suh SG, Gross KC, Byun JK** (1994) *N*-terminal amino acid sequence of persimmon fruit β-galactosidase. Plant Physiol **105:** 975-979

Kim J, Gross KC, Solomos T (1991) Galactose metabolism and ethylene production during development and ripening of tomato fruit. Postharv Biol Technol **1**:67-80

20   **Marck C** (1988) DNA Strider: a "C" program for the fast analysis of DNA and protein sequences on the Apple Macintosh family of computers. Nucleic Acids Res **16**:1829-1836
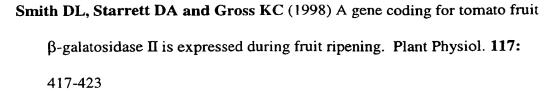
Mitcham EJ, Gross KC, Ng TJ (1989) Tomato fruit cell wall synthesis

during development and senescence. *In vivo* radiolabeling of cell wall

fractions using [^{14}C]sucrose. Plant Physiol **89**:477-481

Nakai K, Kanehisa M (1992) A knowledge base for predicting protein

localization sites in eukaryotic cells. Genomics **14**:897-911

Nielsen H, Engelbrecht J, Brunak S, von Heijne G (1997) Identification of

prokaryotic and eukaryotic signal peptides and prediction of their

cleavage sites. Protein Engineering **10**:1-6

Pressey R (1983) β-Galactosidases in ripening tomatoes. Plant Physiol

71:132-135

Ross GS, Redgwell RJ, MacRae EA (1993) Kiwifruit β-galactosidase:

isolation and activity against specific fruit cell-wall polysaccharides.

Planta **189**:499-506

Ross GS, Wegrzyn T, MacRae EA, Redgwell RJ, (1994) Apple β-

galactosidase. Activity against cell wall polysaccharides and

characterization of a related cDNA clone. Plant Physiol **106**:521-528

Seymour GB, Gross KC (1996) Cell wall disassembly and fruit softening.

Postharvest News Info **7**:45N-52N

Smith CJS, Watson CFS, Ray J, Bird CR, Morris PC, Schuch W,

Grierson D (1988) Antisense RNA inhibition of polygalacturonase gene

expression in transgenic tomatoes. Nature **334**:724-726

Smith DL, Fedoroff NV (1995) LRP1, a gene expressed in lateral and

adventitious root primordia of Arabidopsis. Plant Cell **7**: 735-745

Smith DL, Starrett DA and Gross KC (1998) A gene coding for tomato fruit

β-galatosidase II is expressed during fruit ripening. Plant Physiol. **117:**

417-423

Tieman DM, Harriman RW, Ramamohan G, Handa AK (1992) An

antisense pectin methylesterase gene alters pectin chemistry and soluble

solids in tomato fruit. Plant Cell **4:**667-679

Tigchelaar EC, McGlasson WB, Buescher RW (1978) Genetic regulation of

tomato fruit ripening. HortScience **13:**508-513

Turano FJ, Thakkar SS, Fang T, Weisemann JM (1997) Characterization

and expression of NAD(H) dependent glutamate dehydrogenase genes in

Arabidopsis thaliana. Plant Physiol **113:** 1329-1341

Verwoerd TC, Dekker BMM, Hoekema A (1989) A small-scale procedure

for the rapid isolation of plant RNAs. Nuc Acids Res **17:** 2362

Wegrzyn TF, MacRae EA (1992) Pectinesterase, polygalacturonase, and β-

galactosidase during softening of ethylene-treated kiwifruit. HortScience

27:900-902